

The $\phi = \beta$ Conjecture and Eigenvalues of Random Graph Lifts

Edward Lui

Department of Computer Science
University of British Columbia
lui@cornell.cs.ubc.ca

September 7, 2009

Abstract

Let G be any connected graph, and let λ_1 and ρ denote the spectral radius of G and the universal cover of G , respectively. In [LP09], Linial and Puder have shown that almost every n -lift of G has all of its new eigenvalues bounded by $O(\lambda_1^{1/3} \rho^{2/3})$. Friedman had conjectured that this bound can be improved to $\rho + o_n(1)$ (e.g., see [Fri03, HLW06]).

In [LP09], Linial and Puder have formulated two new categorizations of formal words, namely ϕ and β , which assigns a non-negative integer or infinity to each word. They have shown that for every word w , $\phi(w) = 0$ iff $\beta(w) = 0$, and $\phi(w) = 1$ iff $\beta(w) = 1$. They have conjectured that $\phi(w) = \beta(w)$ for every word w , and they have run extensive numerical simulations that suggest that this conjecture is true. This conjecture, if proven true, gives us a promising approach to proving a slightly weaker version of Friedman's conjecture, namely the bound $O(\rho)$ (see [LP09]).

In this paper, we show that $\phi(w) = 2$ iff $\beta(w) = 2$ for every word w . We also discuss possible strategies for proving $\phi(w) = 3$ iff $\beta(w) = 3$.

Keywords: graph eigenvalues, random graph lifts

1 Introduction

Let G be any connected graph with oriented edges. An n -lift of G is any graph that has an n -fold covering map onto G . Equivalently, an n -lift of G is any graph H with vertices $V(H) = V \times \{1, \dots, n\}$, and whose edges $E(H)$ can

be obtained in the following manner: for every oriented edge $(u, v) \in E(G)$, we choose any permutation $\sigma_{(u,v)} \in S_n$ and add an (undirected) edge between (u, i) and $(v, \sigma_{(u,v)}(i))$ for $i = 1, \dots, n$.

Now, let $E(G) = \{g_1, \dots, g_k\}$. Every k -tuple of permutations in S_n describes an n -lift of G . The random graph model we consider is the probability space $L_n(G)$ of n -lifts of G , with sample space S_n^k and uniform probability distribution. We note that when G is a single vertex with $d/2$ self-loops (with d even), this random graph model is the same as the “permutation model” for random d -regular graphs. For background on lifts and random lifts, see [LR05, AL06, ALM02, HLW06].

Our main interest is to study the eigenvalues of (the adjacency matrix of) random lifts of graphs. Let H be any n -lift of G . The projection $\pi : V(H) \rightarrow V(G)$ defined by $\pi(u, i) = u$ is the natural covering map from H to G . It can be easily verified that if f is an eigenfunction of G , then $f \circ \pi$ is an eigenfunction of H with the same eigenvalue as f . The $|V(G)|$ eigenvalues of H corresponding to the $|V(G)|$ such eigenfunctions are said to be *old*, while the remaining $n|V(G)| - |V(G)|$ eigenvalues of H are said to be *new* (note that a new eigenvalue can have the same value as an old eigenvalue).

Let λ_1 and ρ denote the spectral radius of G and the universal cover of G , respectively. For $H \in L_n(G)$, let $\mu_{\max}(H) = \max\{|\mu| : \mu \text{ is a new eigenvalue of } H\}$. In [Fri03], Friedman showed that almost every n -lift $H \in L_n(G)$ satisfies $\mu_{\max}(H) \leq \lambda_1^{1/2} \rho^{1/2} + o_n(1)$. In [LP09], Linial and Puder improved this bound to $\mu_{\max}(H) \leq \max\left\{1, 3\left(\frac{\rho}{\lambda_1}\right)^{2/3}\right\} \cdot \lambda_1^{1/3} \rho^{2/3} + o_n(1)$ for almost every n -lift $H \in L_n(G)$.

For the special case where G is a single vertex with $d/2$ self-loops (i.e., for the permutation model of random d -regular graphs, with d even), $\mu_{\max}(H)$ corresponds to $\lambda(H) := \max\{|\lambda_2|, |\lambda_n|\}$, where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ are the eigenvalues of H . $\lambda(H)$ has been well-studied in the literature since it gives various information about H , such as the expansion properties of H and the rate of convergence of the random walk on H to the stationary distribution (see [HLW06]).

For the permutation model, Friedman’s result in [Fri03] states that almost every random n -vertex d -regular graph H satisfies $\lambda(H) \leq \sqrt{2d\sqrt{d-1}} + o_n(1)$, which is a slight improvement of the result of Broder and Shamir in [BS87]. Linial and Puder’s result in [LP09] states that almost every random n -vertex d -regular graph H satisfies $\lambda(H) \leq O(d^{2/3})$, and more specifically, $\lambda(H) \leq (4d(d-1))^{1/3} + o_n(1)$ for $d \geq 107$.

For various models of random d -regular graphs (including this specific per-

mutation model), Friedman had shown that almost every random n -vertex d -regular graph H with $d \geq 3$ satisfies $\lambda(H) \leq 2\sqrt{d-1} + o_n(1)$ (see [Fri08]). The Alon-Boppana bound (see [Nil91, Fri03]) shows that $\lambda(H) \geq 2\sqrt{d-1} - o_n(1)$ for every n -vertex d -regular graph H , so Friedman's result cannot be improved significantly, if at all.

The results of [BS87], [Fri03], and [LP09] all use the *Trace Method*, which involves estimating the expected value of the trace of a high power of the adjacency matrix of a random graph. To estimate this expected value, Linial and Puder (in [LP09]) study *word maps* associated with formal words over the alphabet $\Sigma = \Sigma_k = \{g_1^{\pm 1}, \dots, g_k^{\pm 1}\}$. Given a word $w \in \Sigma^*$, the *word map* associated with w maps the k -tuple $(\sigma_1, \dots, \sigma_k) \in S_n^k$ to the permutation $w(\sigma_1, \dots, \sigma_k) \in S_n$, where $w(\sigma_1, \dots, \sigma_k)$ is the permutation obtained by replacing g_1, \dots, g_k with $\sigma_1, \dots, \sigma_k$ (respectively) in the expression for w .

The results of [BS87], [Fri03], and [LP09] all involve studying the probability that 1 (or any given point in $\{1, \dots, n\}$) is a fixed point of the permutation $w(\sigma_1, \dots, \sigma_k)$, when $\sigma_1, \dots, \sigma_k \in S_n$ are chosen randomly with uniform distribution. We are interested in how close this probability is to $\frac{1}{n}$, which depends on the word w .

To study this probability, Linial and Puder (in [LP09]) formulated two new and separate categorizations of formal words, namely $\phi, \beta : \Sigma^* \rightarrow \mathbb{Z}_{\geq 0} \cup \{\infty\}$, which are invariant under reduction of words. Thus, ϕ and β are also categorizations of the words in the free group $\mathbf{F} = \mathbf{F}_k$ generated by $\{g_1, \dots, g_k\}$. Intuitively, $\phi(w)$ measures how close the above probability is to $\frac{1}{n}$ for the word w (the higher $\phi(w)$ is, the closer the probability is to $\frac{1}{n}$). On the other hand, $\beta(w)$ is defined combinatorially without explicit reference to word maps and the symmetric group. Both ϕ and β extend the dichotomy of primitive vs. imprimitive words in \mathbf{F} (recall that $w \in \mathbf{F}$ is said to be *imprimitive* (as an element of \mathbf{F}) if $w = u^d$ for some $u \in \mathbf{F}$ and $d \geq 2$). In some sense, $\phi(w)$ and $\beta(w)$ can be thought of as quantifying the “level of primitivity” of the word w .

In [LP09], Linial and Puder have conjectured that $\phi(w) = \beta(w)$ for every word w . They have proven that $\phi(w) = 0$ iff $\beta(w) = 0$ iff w reduces to the empty word, and $\phi(w) = 1$ iff $\beta(w) = 1$ iff w is imprimitive as an element of \mathbf{F} . These two facts also appear in [BS87] and [Fri03], but not in the explicit language of ϕ and β . Linial and Puder have also made partial progress towards proving $\phi(w) = 2$ iff $\beta(w) = 2$, which allowed them to obtain a slightly better eigenvalue bound compared to the result in [Fri03]. Furthermore, they have run extensive numerical simulations, and the results suggest that $\phi(w) = \beta(w)$ for every word w .

Friedman had conjectured that almost every n -lift $H \in L_n(G)$ of G satisfies

$\mu_{\max}(H) \leq \rho + o_n(1)$ (e.g., see [Fri03, HLW06]). It is known that every n -lift $H \in L_n(G)$ of G satisfies $\mu_{\max}(H) \geq \rho - o_n(1)$ (see [Gre95, Fri03, HLW06]), so one cannot prove a significantly stronger result. The conjecture that $\phi(w) = \beta(w)$, if proven true, gives us a promising approach to proving a slightly weaker version of Friedman's conjecture, namely $\mu_{\max}(H) \leq O(\rho)$ (for almost every n -lift $H \in L_n(G)$) (see [LP09]). Also, if proven true, the conjecture may also significantly simplify the usage of the Trace Method in proving new (or old) eigenvalue bounds in various contexts.

The $\phi(w) = \beta(w)$ conjecture is also interesting in other aspects, such as its connection to word maps. For example, a slightly stronger version of this conjecture made in [LP09] implies that for every word w and sufficiently large n , $w(\sigma_1, \dots, \sigma_k)$ has at least one fixed point on average.

Our work mainly builds on the work of Linial and Puder in [LP09]. In particular, we show that $\phi(w) = 2$ iff $\beta(w) = 2$ for every word $w \in \Sigma^*$. For those already familiar with the paper [LP09], we briefly mention some aspects of our proof. For a word $w \in \Sigma^*$ such that $\beta(w) \geq 2$, recall that there exists a natural surjective function from the (connected) components of some graph Υ to the quotients in \mathcal{Q}_w that have characteristic 2 and type A (see [LP09]). Linial and Puder believed that this function is also injective, which would show that $\phi(w) = 2$ iff $\beta(w) = 2$. In this paper, we follow this strategy and prove that this function is actually injective. This is done by investigating the relationship between the components of Υ and proving a certain property about them. The main technique we use a recursive factoring process that decomposes the word w into finer and finer pieces, which allows us to analyze the components of Υ .

In this paper, we also discuss possible strategies for proving $\phi(w) = 3$ iff $\beta(w) = 3$. Although our result does not give us a new eigenvalue bound, the strategy and techniques we use may be useful in proving the $\phi(w) = \beta(w)$ conjecture, which leads to a promising approach to proving the slightly weaker version of Friedman's conjecture. Our result is also interesting in the context of word maps. For example, the result tells us more about the probability that 1 is a fixed point of the permutation $w(\sigma_1, \dots, \sigma_k)$ (when $\sigma_1, \dots, \sigma_k \in S_n$ are chosen randomly with uniform distribution), and how much this probability differs from $\frac{1}{n}$.

2 Review of ϕ and β and Related Concepts

In this section, we review some concepts and terminology from [LP09], as well as introduce some new terminology for convenience.

Fix $k \geq 1$, and let $\Sigma = \Sigma_k = \{g_1^{\pm 1}, \dots, g_k^{\pm 1}\}$ be an alphabet of $2k$ abstract letters. Σ^* denotes the free monoid of words generated by Σ , and the quotient of Σ^* modulo reduction of words is the free group $\mathbf{F} = \mathbf{F}_k$ generated by $\{g_1, \dots, g_k\}$. For every word $w \in \Sigma^*$ and $n \in \mathbb{Z}^+$, we have a random variable $X_w^{(n)}$ on S_n^k (with the uniform probability distribution); specifically, $X_w^{(n)}(\sigma_1, \dots, \sigma_k)$ is the number of fixed points of $w(\sigma_1, \dots, \sigma_k)$, where $w(\sigma_1, \dots, \sigma_k)$ is the permutation in S_n obtained by replacing g_1, \dots, g_k by $\sigma_1, \dots, \sigma_k$, respectively, in the expression for w . Then, $\Phi_w(n)$ is defined as $\Phi_w(n) = \frac{\mathbb{E}(X_w^{(n)}) - 1}{n} = \frac{\mathbb{E}(X_w^{(n)})}{n} - \frac{1}{n}$. $\frac{\mathbb{E}(X_w^{(n)})}{n}$ is the probability that 1 (or any given point in $\{1, \dots, n\}$) is a fixed point of the permutation $w(\sigma_1, \dots, \sigma_k)$, when $\sigma_1, \dots, \sigma_k \in S_n$ are chosen randomly with uniform distribution. Thus, $\Phi_w(n)$ measures how much this probability differs from $\frac{1}{n}$ for the word w .

It can be shown that for every word $w \in \Sigma^*$, Φ_w can be expressed as a power series in $\frac{1}{n}$; specifically, for every word $w \in \Sigma^*$ and $n \geq |w|$, we have $\Phi_w(n) = \frac{\mathbb{E}(X_w^{(n)}) - 1}{n} = \sum_{i=0}^{\infty} a_i(w) \frac{1}{n^i}$, where the coefficients $a_i(w)$ are integers depending only on w . For any word $w \in \Sigma^*$, $\phi(w)$ is defined to be the least non-negative integer i such that $a_i(w) \neq 0$, or ∞ if no such i exists (i.e., if $\mathbb{E}(X_w^{(n)}) = 1$ for $n \geq |w|$). Thus, $\phi(w)$ measures how much the above probability differs from $\frac{1}{n}$ for the word w . The higher $\phi(w)$ is, the closer the probability is to $\frac{1}{n}$.

We now describe some terminology needed to define $\beta(w)$ for any word $w \in \Sigma^*$. Fix any word $w = g_{i_1}^{\alpha_1} g_{i_2}^{\alpha_2} \dots g_{i_{|w|}}^{\alpha_{|w|}} \in \Sigma^*$, where $i_1, i_2, \dots, i_{|w|} \in \{1, 2, \dots, k\}$ and $\alpha_1, \alpha_2, \dots, \alpha_{|w|} \in \{-1, 1\}$, and consider the following directed edge-labeled graph:

$$\{s_0\} \xrightarrow{g_{i_1}^{\alpha_1}} \{s_1\} \xrightarrow{g_{i_2}^{\alpha_2}} \{s_2\} \xrightarrow{g_{i_3}^{\alpha_3}} \dots \xrightarrow{g_{i_{|w|-1}}^{\alpha_{|w|-1}}} \{s_{|w|-1}\} \xrightarrow{g_{i_{|w|}}^{\alpha_{|w|}}} \{s_{|w|}\}$$

For convenience, we shall change the representation of the edges of the graph so that all the labels are integers in $\{1, \dots, k\}$. Specifically, every edge with label g_i^α is relabeled as i , and if $\alpha = -1$, the orientation of the edge is reversed also. We shall call the resulting graph the *open trail* of w , and we denote it by T_w . E.g., for the word $w = g_1 g_2 g_2 g_2^{-1} g_3 g_2 g_1^{-1}$, the open trail T_w of w is the following directed edge-labeled graph:

$$\{s_0\} \xrightarrow{1} \{s_1\} \xrightarrow{2} \{s_2\} \xrightarrow{2} \{s_3\} \xleftarrow{2} \{s_4\} \xrightarrow{3} \{s_5\} \xrightarrow{2} \{s_6\} \xleftarrow{1} \{s_7\}$$

Fix a word $w \in \Sigma^*$. A *quotient* of T_w is defined as any directed edge-labeled graph whose vertex set is a partition of $\{s_0, \dots, s_{|w|}\}$, and there is

a j -labeled edge (j -edge for short) from vertex U to vertex V iff there exist $U', V' \in V(T_w)$ such that $U' \subseteq U$, $V' \subseteq V$, and T_w contains a j -edge from U' to V' . Equivalently, a quotient of T_w is a graph obtained by “joining” vertices of T_w together in any combination; when two or more vertices are joined together, the resulting vertex is the union of the sets representing the original vertices, and any redundant multiple edges that have the same label and orientation are removed from the resulting graph. All the directed edge-labeled graphs in this paper do not allow two distinct j -edges to have the same head vertex and the same tail vertex.

Given a quotient Γ of T_w , we can also define a quotient of Γ in a similar manner. Specifically, a quotient of Γ is any directed edge-labeled graph Γ' whose vertex set $V(\Gamma')$ is a partition of $\{s_0, \dots, s_{|w|}\}$ that is coarser than $V(\Gamma)$; also, there is a j -edge in Γ' from vertex U' to vertex V' iff there exist $U, V \in V(\Gamma)$ such that $U \subseteq U'$, $V \subseteq V'$, and Γ contains a j -edge from U to V . Again, equivalently, a quotient of Γ is a graph obtained by joining vertices of Γ together in any combination.

We note that a quotient of a quotient of Γ is a quotient of Γ . We also note that there are natural bijections between the following sets: $\{\text{quotients of } \Gamma\}$, $\{\text{partitions of } \{s_0, \dots, s_{|w|}\} \text{ coarser than } V(\Gamma)\}$, and $\{\text{partitions of } V(\Gamma)\}$.

A walk in a quotient of T_w traces out a word in Σ^* . When a j -edge is traversed with the same (respectively, opposite) orientation as the edge, the letter g_j (respectively, g_j^{-1}) is traced out. We note that the word w is traced out by the walk that visits the vertices containing $s_0, \dots, s_{|w|}$, respectively, in this order, and choosing the correct edge to traverse at each step according to w .

Let Γ be any quotient of T_w . Given $s_p, s_q \in \{s_0, \dots, s_{|w|}\}$, we write $s_p \equiv_{\Gamma} s_q$ if s_p and s_q are in the same vertex of Γ , and $s_p \not\equiv_{\Gamma} s_q$ otherwise. Also, given $U, V \subseteq \{s_0, \dots, s_{|w|}\}$, we write $U \equiv_{\Gamma} V$ if $U \subseteq W$ and $V \subseteq W$ for some vertex W in Γ , and $U \not\equiv_{\Gamma} V$ otherwise. We may omit the subscript Γ if the quotient being considered is clear from context.

Let Γ be any quotient of T_w , and let $S = \{\{s_{j_1}, s_{k_1}\}, \dots, \{s_{j_r}, s_{k_r}\}\}$ be any set whose elements are two element subsets of $\{s_0, \dots, s_{|w|}\}$. The *quotient of Γ by S* , denoted Γ/S , is the quotient whose vertex set (a partition of $\{s_0, \dots, s_{|w|}\}$) is defined by the following condition: $s_p, s_q \in \{s_0, \dots, s_{|w|}\}$ are placed in the same block of the partition iff $s_p \equiv s_q$ holds in every quotient of Γ that satisfies $s_{j_i} \equiv s_{k_i}$ for all $i = 1, \dots, r$. One can easily verify that this condition defines a valid partition of $\{s_0, \dots, s_{|w|}\}$ that gives a quotient of Γ . For convenience, we write $\Gamma/\{s_j, s_k\}$ for $\Gamma/\{\{s_j, s_k\}\}$.

Intuitively, for $s_j, s_k \in \{s_0, \dots, s_{|w|}\}$ with $s_j \neq s_k$, $\Gamma/\{s_j, s_k\}$ is the quotient of Γ obtained by joining the two vertices of Γ that contain s_j and s_k ,

respectively. Γ/S is the quotient of Γ obtained by performing this joining operation for each element $\{s_{j_i}, s_{k_i}\}$ in S ; we easily see that the order in which the joining operations are performed does not matter. If S is a set of two element subsets of $V(\Gamma)$, or if S is a partition of $V(\Gamma)$, we define Γ/S in a similar manner.

A quotient of T_w is said to be *realizable* if for every $j \in \{1, \dots, k\}$, the quotient does not contain a pair of distinct j -edges that have the same head vertex or the same tail vertex. Let Γ be any quotient of T_w . The *realization* of Γ , denoted $rel(\Gamma)$, is the quotient of Γ whose vertex set is defined by the following condition: $s_p, s_q \in \{s_0, \dots, s_{|w|}\}$ are placed in the same vertex iff $s_p \equiv s_q$ in every *realizable* quotient of Γ . One can easily verify that this condition defines a valid partition of $\{s_0, \dots, s_{|w|}\}$ that gives a realizable quotient of Γ .

Intuitively, $rel(\Gamma)$ is the quotient of Γ obtained by repeatedly joining any two vertices that “cause the current quotient to be not realizable”, until we get a realizable quotient. If a quotient contains two distinct j -edges that have the same head (respectively, tail) vertex, then the two tail (respectively, head) vertices “cause the quotient to be not realizable”. We easily see that we get the same quotient no matter how we choose exactly which pair of “nonrealizability-causing” vertices to join at each step of the process. It can also be verified that $rel(\Gamma/S)$ can be obtained from Γ by repeatedly joining pairs of vertices that either are specified by S or are causing nonrealizability, in any order, until the quotient is realizable and $s_j \equiv s_k$ for each $\{s_j, s_k\} \in S$.

Given any quotient Γ of T_w , we define the (Euler) *characteristic* of Γ , denoted $\chi(\Gamma)$, to be $\chi(\Gamma) = |E(\Gamma)| - |V(\Gamma)| + 1$. Let S be any set of two element subsets of $\{s_0, \dots, s_{|w|}\}$. The *quotient of T_w generated by S* is the quotient $rel(T_w/S)$. By definition, the quotient generated by S is realizable. Conversely, every realizable quotient Γ of T_w is generated by at least one set of two element subsets of $\{s_0, \dots, s_{|w|}\}$, and any such set is called a *generating set* for Γ . It can be shown that for any realizable quotient Γ of T_w , any minimum generating set (in size) for Γ contains exactly $\chi(\Gamma)$ elements.

The collection of all realizable quotients of T_w such that $s_0 \equiv s_{|w|}$ is denoted by \mathcal{Q}_w . We now have the necessary tools to categorize the quotients in \mathcal{Q}_w and define $\beta(w)$. We say that a quotient $\Gamma \in \mathcal{Q}_w$ *has type A* if one of the smallest generating sets for Γ contains the pair $\{s_0, s_{|w|}\}$; otherwise, we say that Γ *has type B*. Now, we define $\beta(w)$ to be the smallest characteristic of a type-B quotient in \mathcal{Q}_w , or ∞ if no type-B quotient exists.

Earlier, we defined $\Phi_w(n)$ as $\Phi_w(n) = \frac{\mathbb{E}(X_w^{(n)}) - 1}{n}$. However, for $n \geq |w|$, it can be shown that there is an alternative expression for $\Phi_w(n)$ in terms of the

quotients in \mathcal{Q}_w , namely

$$\Phi_w(n) = -\frac{1}{n} + \sum_{\Gamma \in \mathcal{Q}_w} \left(\frac{1}{n}\right)^{\chi(\Gamma)} \frac{\prod_{l=1}^{v_\Gamma-1} (1 - \frac{l}{n})}{\prod_{j=1}^k \prod_{l=1}^{e_\Gamma^j-1} (1 - \frac{l}{n})}, \quad (1)$$

where v_Γ and e_Γ^j denote the number of vertices and the number of j -edges in Γ , respectively.

Let $P(\Gamma)$ denote the big expression inside the sum in (1). Then, $\Phi_w(n)$ can be written more concisely as $\Phi_w(n) = -\frac{1}{n} + \sum_{\Gamma \in \mathcal{Q}_w} P(\Gamma)$. $P(\Gamma)$ is the probability that the walk starting at the point 1 and moving according to the permutation $w(\sigma_1, \dots, \sigma_k)$ traces out the quotient Γ and visits points in the same “coincidence” pattern as the walk in Γ that traces out w and visits the vertices containing $s_0, \dots, s_{|w|}$, respectively, in this order. We note that $P(\Gamma)$ is fully determined by v_Γ and e_Γ^j , $j = 1, \dots, k$, and does not depend on the word w . However, \mathcal{Q}_w does depend on w and is the set of quotients that are summed over in (1). We note that $\sum_{\Gamma \in \mathcal{Q}_w} P(\Gamma) = \frac{\mathbb{E}(X_w^{(n)})}{n}$, which is the probability that 1 is a fixed point of the permutation $w(\sigma_1, \dots, \sigma_k)$.

The power series for Φ_w comes from the expression in (1). Indeed, we can rewrite (1) as

$$\Phi_w(n) = -\frac{1}{n} + \sum_{\Gamma \in \mathcal{Q}_w} \left(\frac{1}{n}\right)^{\chi(\Gamma)} \cdot \prod_{l=1}^{v_\Gamma-1} \left(1 - \frac{l}{n}\right) \cdot \prod_{j=1}^k \prod_{l=1}^{e_\Gamma^j-1} \left(1 + \frac{l}{n} + \frac{l^2}{n^2} + \dots\right) \quad (2)$$

and then further simplify the expression to obtain a power series for Φ_w .

We note that both $\phi(w)$ and $\beta(w)$ are fully determined by the quotients in \mathcal{Q}_w . Thus, we can study both of these parameters by studying \mathcal{Q}_w . In [LP09], Linial and Puder have conjectured that $\phi(w) = \beta(w)$ for every word $w \in \Sigma^*$. They have shown that for every word $w \in \Sigma^*$, $\phi(w) = 0$ iff $\beta(w) = 0$ iff w reduces to the empty word, and $\phi(w) = 1$ iff $\beta(w) = 1$ iff w is imprimitive *as an element of \mathbf{F}* (i.e., $w = u^d$ for some $u \in \mathbf{F}$ and $d \geq 2$). They have also made partial progress on showing that $\phi(w) = 2$ iff $\beta(w) = 2$ for every word $w \in \Sigma^*$.

It is known that both ϕ and β are invariant under reductions, cyclic shifts, cyclic reductions, and automorphisms of \mathbf{F} (i.e. $\phi(f(w)) = \phi(w)$ and $\beta(f(w)) = \beta(w)$ for any word $w \in \Sigma^*$ and any group automorphism f of \mathbf{F} ; $\phi(f(w))$ and $\beta(f(w))$ are well-defined because ϕ and β are invariant under reductions).

The quotients in \mathcal{Q}_w are the realizable quotients of a special graph in \mathcal{Q}_w , namely $rel(T_w/\{s_0, s_{|w|}\})$, which we call the *universal graph of w* and denote by

Γ_w . Recall that any minimum generating set for a realizable quotient Γ of T_w contains exactly $\chi(\Gamma)$ elements. Thus, Γ_w is the only quotient in Q_w that can possibly have characteristic 0, and this occurs iff $\Gamma_w := \text{rel}(T_w/\{s_0, s_{|w|}\}) = \text{rel}(T_w)$, which occurs iff Γ_w has type B; otherwise, Γ_w has type A and characteristic 1. Γ_w is also the only quotient in Q_w that can have characteristic 1 and type A; all other quotients in Q_w with characteristic 1 have type B.

3 The Identification Graph of a Realizable Quotient

In this section, we define the *identification graph* of a realizable quotient, which has already been introduced in [LP09]. However, we discuss the identification graph further and introduce some new terminology. The concepts and terminology here will be used in our proof of $\phi(w) = 2$ iff $\beta(w) = 2$ for every word $w \in \Sigma^*$.

Fix $w \in \Sigma^*$, and let Γ be any realizable quotient of T_w . Suppose that we join two vertices of Γ to obtain another quotient Γ' . Γ' may not be realizable, so additional pairs of vertices may need to be joined in order to obtain the realizable quotient $\text{rel}(\Gamma')$. Thus, joining two vertices of Γ may imply further joins required in order to maintain realizability.

To study exactly which vertices of Γ become joined, we define a new directed edge-labeled graph called the *identification graph* of Γ , denoted I_Γ . The vertices of I_Γ are pairs of distinct vertices of Γ , and there is a j -edge from vertex $\{U, V\}$ to vertex $\{U', V'\}$ iff Γ contains a j -edge from U to some vertex in $\{U', V'\}$, and a j -edge from V to the other vertex in $\{U', V'\}$ (since Γ is realizable, these two j -edges cannot point to the same vertex in $\{U', V'\}$). We note that for every pair $\{e, e'\}$ of distinct j -edges in Γ , there is a j -edge in I_Γ from $\{\text{tail}(e), \text{tail}(e')\}$ to $\{\text{head}(e), \text{head}(e')\}$ (where $\text{head}(e)$ refers to the vertex pointed to by the directed edge e). These are all the edges of I_Γ , so I_Γ contains exactly $\sum_{j=1}^k \binom{e_\Gamma^j}{2}$ edges. We also note that I_Γ does not contain a pair of distinct j -edges that have the same head vertex or the same tail vertex, since Γ is realizable. For convenience, we shall write I_w for I_{Γ_w} .

Given any directed edge-labeled graph G and a vertex $v \in G$, let $d_G^{j+}(v) = |\{e \in E(G) : \text{tail}(e) = v\}|$, $d_G^{j-}(v) = |\{e \in E(G) : \text{head}(e) = v\}|$, and $d_G^j(v) = d_G^{j+}(v) + d_G^{j-}(v)$. Subscripts are omitted when the graph under consideration is clear from context.

Proposition 1. *Let $w \in \Sigma^*$, and let Γ be any realizable quotient of T_w . Then, for every $\{U, V\} \in I_\Gamma$, we have $d(\{U, V\}) \leq \min\{d(U), d(V)\}$. In particular,*

we have $\delta(I_\Gamma) \leq \delta(\Gamma)$ (provided that $V(I_\Gamma)$ is nonempty).

Proof. Let $\{U, V\} \in I_\Gamma$. WLOG, suppose that $d(U) = \min\{d(U), d(V)\}$. Let $j \in \{1, \dots, k\}$. It suffices to show that $d^{j+}(\{U, V\}) \leq d^{j+}(U)$ and $d^{j-}(\{U, V\}) \leq d^{j-}(U)$. Since I_Γ does not contain a pair of distinct j -edges that have the same head vertex or the same tail vertex, we have $d^{j+}(\{U, V\}) \leq 1$ and $d^{j-}(\{U, V\}) \leq 1$. Now we note that if $d^{j+}(\{U, V\}) = 1$, then by the construction of I_Γ , there exists a j -edge in Γ from U to some vertex, so $d^{j+}(U) \geq 1$. Similarly, if $d^{j-}(\{U, V\}) = 1$, then $d^{j-}(U) \geq 1$. \square

Let $w \in \Sigma^*$. Since both ϕ and β are invariant under cyclic reductions, we can assume that w is cyclically reduced. Furthermore, we already know exactly when $\phi(w) = 0$, $\beta(w) = 0$, $\phi(w) = 1$, and $\beta(w) = 1$. Thus, in our investigation of $\phi(w)$ and $\beta(w)$, we can also assume that $\phi(w) \geq 2$ and $\beta(w) \geq 2$. We will often be making these convenient assumptions.

Proposition 2. *Let $w \in \Sigma^*$ be a cyclically reduced word such that $\phi(w) \geq 2$ (and hence $\beta(w) \geq 2$). Then, I_w is a disjoint union of paths (including isolated vertices) with some orientation of its edges (not necessarily the same orientation within each path). In particular, I_w contains no cycles (with any orientation of each of the cycle's edges).*

Proof. Since w is cyclically reduced and is not the empty word (since $\phi(w) \geq 2$), Γ_w is a cycle. Thus, $\delta(\Gamma_w) = 2$, so $\delta(I_w) \leq 2$ by Proposition 1. Thus, I_w is a disjoint union of paths and cycles. It has already been shown in [LP09] that I_w contains no cycles. \square

Let $w \in \Sigma^*$, and let Γ be a realizable quotient of T_w . Using I_Γ , we can see what additional vertices are joined when pairs of vertices are joined in Γ and then a realization operation is performed. For example, suppose that distinct vertices v_1, v_2 are joined in Γ , and then a realization operation is performed to obtain the quotient $rel(\Gamma/\{v_1, v_2\})$. Let C be the component of I_Γ containing the vertex $\{v_1, v_2\}$. By the construction of I_Γ , all the pairs of vertices of Γ in the component C also become joined during the realization operation. It is clear that joining any of the pairs of vertices of Γ in C would yield the same quotient after the realization operation is performed. Not only do the pairs in C become joined, but the pairs in other components of I_Γ may become joined as well. For example, suppose that $\{v_2, v_3\}$ is also in the component C . Then, if v_1 and v_2 are joined in Γ , v_2 and v_3 also become joined. By transitivity, these two joins automatically imply that v_1 and v_3 are joined. However, $\{v_1, v_3\}$ may not be in the component C , but instead may be in the component C' . Thus, all the pairs in C' would be joined as well.

To determine exactly which pairs of vertices of Γ become joined, we start with the set S consisting of all the vertices in C . Then, if S contains $\{v_i, v_j\}$ and $\{v_j, v_k\}$ with $v_i \neq v_k$ but does not contain $\{v_i, v_k\}$, we take all the vertices in the component of I_Γ that contains $\{v_i, v_k\}$, and we add these vertices to S . We repeat the previous step until no more vertices need to be added to S . We note that S naturally defines a symmetric relation on $V(\Gamma)$, and we can insist that the relation defined is also reflexive. By our construction of S , the relation defined by S is also transitive, so S defines an equivalence relation on $V(\Gamma)$. We note that S also induces a set of components of I_Γ . We see that S describes exactly which pairs of vertices of Γ become joined; i.e., the set S contains precisely the pairs $\{v, v'\} \in V(I_\Gamma)$ for which $v \equiv v'$ in $rel(\Gamma/\{v_1, v_2\})$.

We note that if T is a set of vertices of I_Γ that defines a reflexive, symmetric relation on $V(\Gamma)$ that is also transitive (i.e., an equivalence relation), and T induces a set of components of I_Γ , then T naturally defines a realizable quotient of Γ . The equivalence relation defined by T gives a valid quotient of Γ , and the condition that T induces a set of components of I_Γ ensures that the quotient is realizable.

Let Y denote the set of components of I_Γ induced by S . We see that Y is the (unique) smallest set of components of I_Γ that contains C and is *closed under transitivity of pairs*, i.e., if the collection of components contains the pairs $\{v_i, v_j\}$ and $\{v_j, v_k\}$ with $v_i \neq v_k$, then the set of components also contains the pair $\{v_i, v_k\}$. (In other words, a collection of components is *closed under transitivity of pairs* if the reflexive, symmetric relation on $V(\Gamma)$ defined by the set of vertices in the collection of components is transitive.) We easily see that if Z is another set of components of I_Γ that contains C and is closed under transitivity of pairs, then $Y \subseteq Z$.

We say that a component C of I_Γ *implies* component C' if for every (equivalently, some) pair $\{U, V\} \in C'$, we have $U \equiv V$ in $rel(\Gamma/V(C))$. Intuitively, C implies C' if joining a pair of vertices of Γ in C implies that each of the pairs of vertices of Γ in C' also become joined when a realization operation is performed. From our discussion above, we see that the set of components implied by a component C is the smallest set of components that contains C and is closed under transitivity of pairs. Furthermore, the set of components implied by C is contained in any such set of components.

4 Proof of $\phi(w) = 2$ iff $\beta(w) = 2$

The main result of this section is that $\phi(w) = 2$ iff $\beta(w) = 2$ for every word $w \in \Sigma^*$. In the proof, we will be factoring the word w recursively into finer

and finer pieces. Lemma 2 below helps us perform this factoring process, while Lemma 1 is used to prove Lemma 2. We first prove Lemma 1, which is related to factoring certain words into a particular form. Let ϵ denote the empty word.

Lemma 1. *Let $x, y, z \in \Sigma^*$ with $xy = yz$ and $x \neq \epsilon$. Then $x = uv, y = (uv)^p u$, and $z = vu$ for some $u, v \in \Sigma^*$ and $p \geq 0$.*

Proof. We prove the lemma by induction on $|y|$. If $|y| = 0$ (i.e., $y = \epsilon$), then $x = z$, so we let $u = \epsilon, v = x$, and $p = 0$, as required. Now, suppose that $|y| \geq 1$.

If $|y| \leq |x|$, then since $xy = yz$, y is a prefix of x . Thus, $x = yx'$ for some $x' \in \Sigma^*$. Since $xy = yz$, we have $yx'y = yz$, so $x'y = z$. Now, we let $u = y, v = x'$, and $p = 0$, as required.

If $|y| > |x|$, then since $xy = yz$, x is a prefix of y . Thus, $y = xy'$ for some $y' \in \Sigma^*$. Since $x \neq \epsilon$, we have $|y'| < |y|$. $xy = yz$ implies that $xx'y' = xy'z$, so $xy' = y'z$. By the induction hypothesis, we have $x = uv, y' = (uv)^p, z = vu$ for some $u, v \in \Sigma^*$ and $p \geq 0$. We also have $y = xy' = uv(uv)^p = (uv)^{p+1}$, as required. \square

We now prove Lemma 2, which allows us to factor a pair of words with differing last letter into one of several different forms. Let $\Sigma^+ = \Sigma^* \setminus \{\epsilon\}$, and for $x \in \Sigma^+$, let $x_{|x|}$ denote the last letter of x .

Lemma 2. *Let $x, y \in \Sigma^+$ such that $x_{|x|} \neq y_{|y|}$. Then, x and y can be factored into one of the following forms:*

- (1) $x = uv$ and $y = (uv)^p u b$ for some $u, v, b \in \Sigma^+$ and $p \geq 0$ such that $b_1 \neq v_1$ and $b_{|b|} \neq v_{|v|}$.
- (2) $x = u$ and $y = u^p b$ for some $u, b \in \Sigma^+$ and $p \geq 0$ such that $b_1 \neq u_1$ and $b_{|b|} \neq u_{|u|}$.
- (3) $x = uv$ and $y = (uv)^p u$ for some $u, v \in \Sigma^+$ and $p \geq 0$ such that $u_{|u|} \neq v_{|v|}$.
- (4) $x = u$ and $y = u^p$ for some $u \in \Sigma^+$ and $p \geq 1$.

Proof. Let e be the longest common prefix of xy and y . Then, $y = eb$ for some $b \in \Sigma^*$, so $xy = xeb$. We can also write $xy = ecb$ for some $c \in \Sigma^+$. Thus, $xeb = ecb$, so $xe = ec$. By Lemma 1, we have $x = uv, e = (uv)^p u$, and $c = vu$ for some $u, v \in \Sigma^*$ and $p \geq 0$. Thus, $y = eb = (uv)^p u b$. Since $x \neq \epsilon$, we cannot have $u = v = \epsilon$.

If $u \neq \epsilon$, $v \neq \epsilon$, and $b \neq \epsilon$, we have $b_1 \neq c_1$ by definition of e , so $b_1 \neq v_1$. Since $x_{|x|} \neq y_{|y|}$, we have $b_{|b|} \neq v_{|v|}$. Thus, we can write x and y in the form (1).

If $b \neq \epsilon$ and exactly one of u, v is equal to ϵ , then we have $b_1 \neq c_1$ by definition of e . Letting z equal the word in $\{u, v\}$ that is nonempty, we see that $b_1 \neq z_1$ and $b_{|b|} \neq z_{|z|}$, as above. Renaming z as u , we can write x and y in the form (2).

If $u \neq \epsilon$, $v \neq \epsilon$, and $b = \epsilon$, then we have $u_{|u|} \neq v_{|v|}$, since $x_{|x|} \neq y_{|y|}$. Thus, we can write x and y in the form (3).

If $b = \epsilon$ and exactly one of u, v is equal to ϵ , we can write x and y in the form (4) (by renaming the nonempty word in $\{u, v\}$ as u). We note that $p \geq 1$ in this case because $y \neq \epsilon$. \square

Theorem 1. *Let w be any word in Σ^* . Then, $\phi(w) = 2$ iff $\beta(w) = 2$.*

Proof. WLOG, we can assume that w is cyclically reduced, $\phi(w) \geq 2$, and $\beta(w) \geq 2$. Since $\beta(w) \geq 2$, no quotients in \mathcal{Q}_w have characteristic 0, and only the universal graph Γ_w has characteristic 1. Thus, one can easily verify from (2) that $a_2(w) = -\binom{v_{\Gamma_w}}{2} + \sum_{j=1}^k \binom{e_{\Gamma_w}^j}{2} + |\{\Gamma \in \mathcal{Q}_w : \chi(\Gamma) = 2\}|$. We will show that $|\{\Gamma \in \mathcal{Q}_w : \chi(\Gamma) = 2 \text{ and } \Gamma \text{ has type A}\}| = \binom{v_{\Gamma_w}}{2} - \sum_{j=1}^k \binom{e_{\Gamma_w}^j}{2}$. It then follows that $a_2(w) = |\{\Gamma \in \mathcal{Q}_w : \chi(\Gamma) = 2 \text{ and } \Gamma \text{ has type B}\}|$, so $\phi(w) = 2$ iff $\beta(w) = 2$.

To show that $|\{\Gamma \in \mathcal{Q}_w : \chi(\Gamma) = 2 \text{ and } \Gamma \text{ has type A}\}| = \binom{v_{\Gamma_w}}{2} - \sum_{j=1}^k \binom{e_{\Gamma_w}^j}{2}$, we first note that $\binom{v_{\Gamma_w}}{2} - \sum_{j=1}^k \binom{e_{\Gamma_w}^j}{2}$ is equal to the number of (connected) components of the identification graph I_w , since I_w contains no cycles (Proposition 2). It suffices to show that the number of components of I_w is equal to $|\{\Gamma \in \mathcal{Q}_w : \chi(\Gamma) = 2 \text{ and } \Gamma \text{ has type A}\}|$. Each component C of I_w gives a quotient in \mathcal{Q}_w that has characteristic 2 and type A, namely $\text{rel}(\Gamma_w/V(C))$, or equivalently $\text{rel}(\Gamma_w/\{U, V\})$ for any $\{U, V\} \in V(C)$. (Each quotient $\text{rel}(\Gamma_w/V(C))$ cannot have characteristic less than 2 because if it did, it would have type B.) This defines a surjective function f from the components of I_w to the quotients in \mathcal{Q}_w that have characteristic 2 and type A. It suffices to show that f is also injective.

Firstly, we recall that the components of I_w are paths (Proposition 2). Thus, each component has a length (isolated vertices have length 0). Recall that we say component C implies component C' if $U \equiv_{f(C)} V$ for every pair $\{U, V\} \in C'$. We will prove the following claim: If a component C implies a different component C' , then the length of C is strictly longer than the length of C' . This proves the main theorem, since if $f(C) = f(C')$, then C and C'

would imply one another, and if $C \neq C'$, then C would be strictly longer than C' , and C' would be strictly longer than C ; this is clearly a contradiction.

We now prove the above claim. Let C be any component of I_w , and let $\{U_0, V_0\}$ and $\{U_l, V_l\}$ be the two endpoints of C . The non-backtracking walk from $\{U_0, V_0\}$ to $\{U_l, V_l\}$ corresponds to a pair of distinct non-backtracking walks in Γ_w . (These two walks are non-backtracking because I_w does not contain a pair of distinct j -edges that have the same head vertex or the same tail vertex). WLOG, we can assume that the two walks are from U_0 to U_l and from V_0 to V_l ; we denote the two walks by w_{U_0, U_l} and w_{V_0, V_l} , respectively.

We claim we can assume that the two subgraphs traced out by the two walks do intersect. To see this, suppose that the two subgraphs don't intersect. Then, none of the vertices in C intersect (as sets) nontrivially with one another. Thus, C is closed under transitivity of pairs, so C does not imply any components other than itself, and we are done. (Recall that the set of components implied by C is the smallest set of components that contains C and is closed under transitivity of pairs.)

Since w is nonempty and cyclically reduced, Γ_w is a cycle. We now claim that the two walks must move in the same direction along the cycle. Suppose not. We first note that the two walks start at different vertices. Since the two subgraphs traced out by the two walks do intersect, the two walks must meet each other either by visiting the same vertex simultaneously, or by traversing the same edge simultaneously. By the construction of I_w , both walks simultaneously traverse same-labeled edges with the same orientation (either both consistent with the edges' orientation, or both inconsistent). Thus, visiting the same vertex simultaneously would imply that two distinct j -edges have the same head or the same tail, contradicting the realizability of Γ_w . Traversing the same edge simultaneously is clearly impossible.

We recall that the set of components implied by C is contained in any set of components (of I_w) that contains C and is closed under transitivity of pairs. Thus, to complete the proof, it suffices to describe a set of components that is closed under transitivity of pairs and whose unique longest component is C . Our strategy is to recursively factor the word w into finer and finer pieces (using Lemma 2), until certain conditions are met. At the end of the factoring process, we will have a factorization of w that allows us to describe a set of components that have the desired properties.

We now describe how to obtain our initial factorization of w . Since I_w contains no cycles, the two walks w_{U_0, U_l} and w_{V_0, V_l} cannot go around the cycle of Γ_w and return back to their respective starting vertices. WLOG, we can assume that $s_0 \in U_0$ and that w_{U_0, U_l} traces out a prefix of w ; i.e., the word w is traced out by the walk that starts at U_0 , moves in the same direction as w_{U_0, U_l} ,

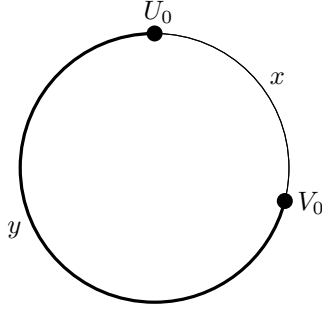


Figure 1: Example of how Γ_w may look like. The two walks w_{U_0, U_l} and w_{V_0, V_l} start at vertices U_0 and V_0 , respectively, and move clockwise around the cycle for some distance. The word w is traced out by the walk from U_0 to itself, going clockwise around the cycle once. x is the prefix of w traced out by the walk from U_0 to V_0 , and y is the remaining suffix of w .

and goes once around the cycle back to U_0 . This assumption is possible because if w' is a cyclic shift of w , or if $w' = w^{-1}$, then Γ_w and $\Gamma_{w'}$ are isomorphic as directed edge-labeled graphs; thus, there is a natural isomorphism between I_w and $I_{w'}$ that preserves the implications of each component.

Let x be the prefix of w traced out by the walk from U_0 to V_0 , and let y be the remaining suffix of w so that $w = xy$ (see Figure 1). We note that $x \neq \epsilon$ and $y \neq \epsilon$. Also, since $\{U_0, V_0\}$ is an endpoint of C , we have $x_{|x|} \neq y_{|y|}$. Using Lemma 2, we factor w into the form $w^{(1)}$, where $w^{(1)}$ is one of the following expressions:

- (1) $w^{(1)} = (u^{(1)}v^{(1)})^{p_1}u^{(1)}b^{(1)}$ where $p_1 \geq 1$, $b_1^{(1)} \neq v_1^{(1)}$, and $b_{|b|}^{(1)} \neq v_{|v|}^{(1)}$.
- (2) $w^{(1)} = (u^{(1)})^{p_1}b^{(1)}$ where $p_1 \geq 2$, $b_1^{(1)} \neq u_1^{(1)}$, and $b_{|b|}^{(1)} \neq u_{|u|}^{(1)}$.
- (3) $w^{(1)} = (u^{(1)}v^{(1)})^{p_1}u^{(1)}$ where $p_1 \geq 1$ and $u_{|u|}^{(1)} \neq v_{|v|}^{(1)}$.
- (4) $w^{(1)} = (u^{(1)})^{p_1}$ where $p_1 \geq 2$.

Above, $u^{(1)}$, $v^{(1)}$, and $b^{(1)}$ are elements of Σ^+ . In (2) above, we have $p_1 \geq 2$ because $p \geq 1$ in (2) of Lemma 2, since $p = 0$ would imply that $y = b$, so $y_1 \neq x_1$, contradicting the fact that w_{U_0, U_l} and w_{V_0, V_l} intersect. We note that w cannot be factored into the form (4) above, since w is not imprimitive (since $\phi(w) \neq 1$). Figure 2 shows how the factorization $w^{(1)}$ of w may look like in the universal graph Γ_w .

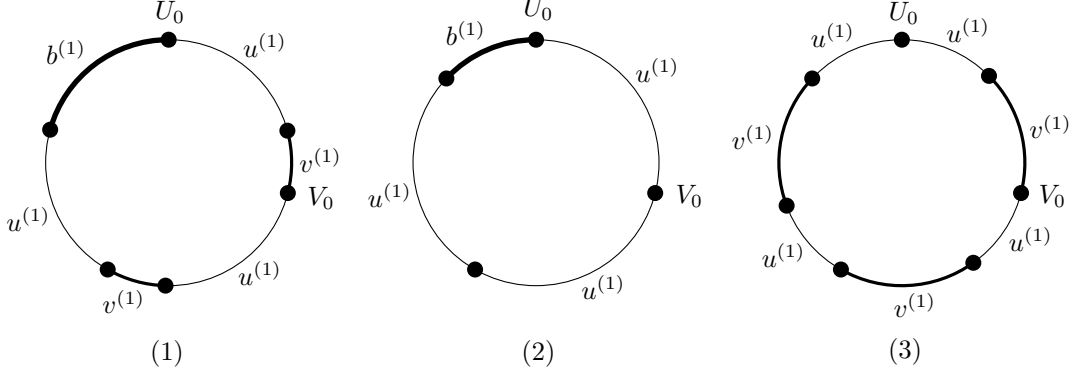


Figure 2: Examples of how the factorization $w^{(1)}$ of w may look like in Γ_w . The three graphs correspond to the factorizations (1) – (3) above. Note that factorization (4) is not possible because w is not imprimitive.

We will later describe how to further factor w , if necessary. As we recursively factor w , the expressions for w become increasingly complicated, so we will not be writing them out explicitly. However, the first factorization of w has certain properties, and these properties are preserved throughout the factoring process. The stopping condition for the factoring process guarantees that the final factorization has certain additional properties. All of these properties combined together allow us to describe a set of components that is closed under transitivity of pairs and whose unique longest component is C .

To help us describe such a set of components, we introduce the notion of a “path” in an explicit factorization of w . “Paths” in the final factorization of w will be used to describe a set of components that is closed under transitivity of pairs. To show that C is the longest component in this set of components, we will describe a property regarding “paths” that the factorization $w^{(1)}$ has, and we will show that this property is preserved throughout the factoring process. In the final factorization of w , we will be able to use this property to show that C is the longest component in the set of components.

Given an explicit expression $w^{(i)}$ representing a factorization of w , we define a *path* in $w^{(i)}$ to be a 2-tuple (n, z) where n is an integer specifying the position of the symbol the path starts at, and z is a finite subexpression and prefix of $s(w^{(i)})^\infty$ ($(w^{(i)})^\infty$ is the concatenation of $w^{(i)}$ with itself infinitely many times), where s is the subexpression and suffix of $w^{(i)}$ starting at the n th symbol. For example, if $w^{(i)} = uvzxyzzy$, then $(2, vzx)$, $(7, zyuvz)$, and $(4, xyzzyuvzxy)$ are all paths in $w^{(i)}$. Note that in the definition of a path in $w^{(i)}$, the symbols in $w^{(i)}$ are regarded as formal symbols, so the actual words in Σ^* represented

by the symbols do not matter.

The two subexpressions for two paths in $w^{(i)}$ are equivalent iff both subexpressions contain the exact same sequence of formal symbols, in which case we say that the two paths are *subexpression-equivalent*. A path is also said to *trace out* the word (in Σ^*) represented by the path's subexpression. A pair of subexpression-equivalent paths is said to be of *maximal length* if the two paths cannot be extended (in the forward direction, while keeping the starting points fixed); i.e., the symbol following the end of one path is not the same as the symbol following the end of the other path.

Let m denote the position of the letter in w (with all of its letters written out) corresponding to the first letter of the second occurrence of $u^{(1)}$ in $w^{(1)}$. We note that the pair of walks in Γ_w corresponding to C starts at the pair of vertices right before w_1 and w_m in Γ_w . m remains fixed throughout the rest of the proof.

One can easily verify that $w^{(1)}$ satisfies properties (5) – (7) below (with $i = 1$; properties (5a) – (5c) correspond to (1) – (3) above):

(5) $w^{(i)}$ satisfies one of the following properties:

- (5a) $w^{(i)}$ is some expression in $u^{(i)}, v^{(i)}, b^{(i)} \in \Sigma^+$ (using all of these symbols and nothing else) such that $b_1^{(i)} \neq v_1^{(i)}$ and $b_{|b^{(i)}|}^{(i)} \neq v_{|v^{(i)}|}^{(i)}$. Also, no occurrence of $u^{(i)}$ in $w^{(i)}$ is cyclically followed or preceded by $u^{(i)}$. (By “cyclically”, we mean that the symbol following the last symbol of $w^{(i)}$ is the first symbol of $w^{(i)}$, and the symbol preceding the first symbol is the last symbol.)
- (5b) $w^{(i)}$ is some expression in $u^{(i)}, b^{(i)} \in \Sigma^+$ (using all of these symbols and nothing else) such that $b_1^{(i)} \neq u_1^{(i)}$ and $b_{|b^{(i)}|}^{(i)} \neq u_{|u^{(i)}|}^{(i)}$.
- (5c) $w^{(i)}$ is some expression in $u^{(i)}, v^{(i)} \in \Sigma^+$ (using all of these symbols and nothing else) such that $u_{|u^{(i)}|}^{(i)} \neq v_{|v^{(i)}|}^{(i)}$.

(6) Any occurrence of $v^{(i)}$ or $b^{(i)}$ in $w^{(i)}$ is cyclically followed and preceded by $u^{(i)}$.

(7) The first symbol of $w^{(i)}$ is $u^{(i)}$. The letter w_m corresponds to the first letter of a non-first occurrence of $u^{(i)}$ in $w^{(i)}$. The pair of distinct subexpression-equivalent paths in $w^{(i)}$ that traces out the longest word is the pair of paths that start at the first occurrence of $u^{(i)}$ and the occurrence of $u^{(i)}$ described above, respectively, and is of maximal length. Furthermore, this pair of paths is unique.

Now, we describe a process that generates a finite sequence $w^{(1)}, \dots, w^{(N)}$ of factorizations of w . We will show that each of the factorizations satisfies properties (5) – (7). We start with the expression $w^{(1)}$, which satisfies properties (5) – (7). Let $i \geq 1$, and suppose that $w^{(i)}$ satisfies properties (5) – (7). If $w^{(i)}$ satisfies (5a) or (5b), we stop; otherwise, $w^{(i)}$ must satisfy (5c), i.e. $w^{(i)}$ is some expression in $u^{(i)}, v^{(i)} \in \Sigma^+$ such that $u_{|u^{(i)}|}^{(i)} \neq v_{|v^{(i)}|}^{(i)}$. We will show how to generate the next factorization $w^{(i+1)}$ from $w^{(i)}$.

Using Lemma 2, we explicitly factor $u^{(i)}$ and $v^{(i)}$ into one of the following forms:

- (8) $u^{(i)} = u^{(i+1)}v^{(i+1)}$ and $v^{(i)} = (u^{(i+1)}v^{(i+1)})^{p_{i+1}}u^{(i+1)}b^{(i+1)}$ where $p_{i+1} \geq 0$, $b_1^{(i+1)} \neq v_1^{(i+1)}$, and $b_{|b^{(i+1)}|}^{(i+1)} \neq v_{|v^{(i+1)}|}^{(i+1)}$.
- (9) $u^{(i)} = u^{(i+1)}$ and $v^{(i)} = (u^{(i+1)})^{p_{i+1}}b^{(i+1)}$ where $p_{i+1} \geq 0$, $b_1^{(i+1)} \neq u_1^{(i+1)}$, and $b_{|b^{(i+1)}|}^{(i+1)} \neq u_{|u^{(i+1)}|}^{(i+1)}$.
- (10) $u^{(i)} = u^{(i+1)}v^{(i+1)}$ and $v^{(i)} = (u^{(i+1)}v^{(i+1)})^{p_{i+1}}u^{(i+1)}$ where $p_{i+1} \geq 0$ and $u_{|u^{(i+1)}|}^{(i+1)} \neq v_{|v^{(i+1)}|}^{(i+1)}$.
- (11) $u^{(i)} = u^{(i+1)}$ and $v^{(i)} = (u^{(i+1)})^{p_{i+1}}$ where $p_{i+1} \geq 1$.

Above, $u^{(i+1)}$, $v^{(i+1)}$, and $b^{(i+1)}$ are words in Σ^+ . We note that $u^{(i)}$ and $v^{(i)}$ cannot be factored into the form (11), since this would imply that w is imprimitive, which is a contradiction. To obtain the factorization $w^{(i+1)}$ of w , we substitute $u^{(i)}$ and $v^{(i)}$ in $w^{(i)}$ by the appropriate expressions above.

For example, if $w^{(1)}$ satisfies (5a) or (5b) (i.e., $w^{(1)}$ is equal to the expression in (1) or (2) shown earlier), we stop; otherwise, $w^{(1)}$ must satisfy (5c), i.e., $w^{(1)}$ is equal to the expression in (3) shown earlier. Thus, $w^{(1)} = (u^{(1)}v^{(1)})^{p_1}u^{(1)}$. Depending on whether we have factored $u^{(1)}$ and $v^{(1)}$ into the form (8), (9), or (10), we have $w^{(2)} = ((u^{(2)}v^{(2)})^{p_2+1}u^{(2)}b^{(2)})^{p_1}u^{(2)}v^{(2)}$, $w^{(2)} = ((u^{(2)})^{p_2+1}b^{(2)})^{p_1}u^{(2)}$, or $w^{(2)} = ((u^{(2)}v^{(2)})^{p_2+1}u^{(2)})^{p_1}u^{(2)}v^{(2)}$, respectively. Figure 3 shows how the factorization $w^{(2)}$ of w may look like in the universal graph Γ_w .

Claim 1. $w^{(i+1)}$ satisfies properties (5) – (7).

Proof of Claim 1. By comparing the expressions in (8) – (10) to properties (5a) – (5c) respectively, we see that $w^{(i+1)}$ satisfies property (5). Using the fact that $w^{(i)}$ satisfies properties (5c) and (6), one can easily verify (by considering cases (8) – (10) separately) that $w^{(i+1)}$ also satisfies property (6).

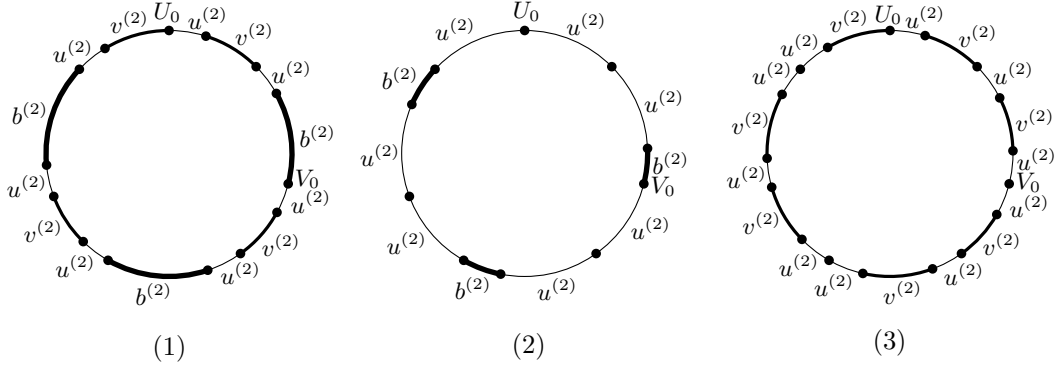


Figure 3: Examples of how the factorization $w^{(2)}$ of w may look like in Γ_w .

Since $w^{(i)}$ satisfies property (7) and we factored $u^{(i)}$ into some expression that starts with $u^{(i+1)}$, the first symbol of $w^{(i+1)}$ is $u^{(i+1)}$, and the letter w_m corresponds to the first letter of a non-first occurrence of $u^{(i+1)}$ in $w^{(i+1)}$. Now, we shall determine which pairs of distinct subexpression-equivalent paths in $w^{(i+1)}$ trace out the longest word, if such a pair exists.

We note that such a pair must exist, since any pair of distinct subexpression-equivalent paths in $w^{(i+1)}$ cannot trace a word that is at least $|w|$ in length, since this would imply that I_w contains a cycle, which is a contradiction. Now, consider any pair of distinct subexpression-equivalent paths in $w^{(i+1)}$. If the pair of paths starts with $v^{(i+1)}$ or $b^{(i+1)}$, then there is another pair of paths that trace out a longer word, since both $v^{(i+1)}$ and $b^{(i+1)}$ are always cyclically preceded by $u^{(i+1)}$. Thus, we can assume that the pair of paths starts with $u^{(i+1)}$.

Some of the $u^{(i+1)}$'s in $w^{(i+1)}$ correspond to the $u^{(i)}$'s in $w^{(i)}$ in the sense that the $u^{(i+1)}$'s come from the expression we wrote for $u^{(i)}$ (as opposed to coming from the expression we wrote for $v^{(i)}$). Such $u^{(i)}$'s are called old, while the other $u^{(i)}$'s are called new.

Consider the pair of $u^{(i+1)}$'s that the pair of paths start with. If the pair of initial $u^{(i+1)}$'s are both new, then there is another pair of paths that trace out a longer word, since both $u^{(i+1)}$'s are cyclically preceded by the same symbol in $w^{(i+1)}$. (Consider cases (8) – (10) separately, and recall that $v^{(i)}$ is always cyclically preceded by $u^{(i)}$ in $w^{(i)}$.)

If one of the initial $u^{(i+1)}$'s is old while the other is new, then one can easily verify that the word traced out by the pair of paths is at most $|v^{(i)}|$ in length; let M be the smallest nonnegative integer that is \geq the length of the word traced out by any such pair of paths in $w^{(i+1)}$. (For the case where we

have written $u^{(i)}$ and $v^{(i)}$ in the form (10), recall that $v^{(i)}$ is always cyclically followed by $u^{(i)}$ in $w^{(i)}$.) One can easily verify that for case (8), we have $M = |(u^{(i+1)}v^{(i+1)})^{p_{i+1}}u^{(i+1)}|$; for case (9), we have $M = |(u^{(i+1)})^{p_{i+1}}b^{(i+1)}|$; and for case (10), we have $M = |v^{(i)}| = |(u^{(i+1)}v^{(i+1)})^{p_{i+1}}u^{(i+1)}|$.

Now, suppose that both of the initial $u^{(i+1)}$'s are old. Then, the pair of initial $u^{(i+1)}$'s corresponds to a pair of $u^{(i)}$'s in $w^{(i)}$. Let A denote the set of pairs of distinct subexpression-equivalent paths in $w^{(i)}$ that both start with $u^{(i)}$ and are of maximal length. Let B denote the set of pairs of distinct subexpression-equivalent paths in $w^{(i+1)}$ that both start with old $u^{(i+1)}$ and are of maximal length. We note that there is a natural bijection between A to B . Specifically, a pair of paths in B that both start with old $u^{(i+1)}$ corresponds to the pair of paths in A that start at the corresponding pair of $u^{(i)}$'s in $w^{(i)}$.

We claim that every pair of paths in B traces out a word that is exactly M letters (in Σ) longer than the corresponding pair of paths in A . To see this, consider any pair of paths in A ; one of the paths must end right before a $u^{(i)}$ while the other path must end right before a $v^{(i)}$ (since $w^{(i)}$ is an expression in $u^{(i)}$ and $v^{(i)}$, and the pair is of maximal length). We note that the corresponding pair of paths in B would reach the expressions used to substitute this specific occurrence of $u^{(i)}$ and $v^{(i)}$. Then, looking at the expressions we wrote for $u^{(i)}$ and $v^{(i)}$, we see that the claim holds.

Thus, a pair of paths in B that traces out the longest word (out of all the pairs of paths in B) corresponds to a pair of paths in A that traces out the longest word (out of all the pairs of paths in A). Since $w^{(i)}$ satisfies property (7), we now see that $w^{(i+1)}$ also satisfies property (7). \square

If we have factored $u^{(i)}$ and $v^{(i)}$ into the form (8) or (9), then $w^{(i+1)}$ satisfies property (5a) or (5b), so we stop. Otherwise, we must have factored $u^{(i)}$ and $v^{(i)}$ into the form (10). In this case, we have $\max\{|u^{(i+1)}|, |v^{(i+1)}|\} < |u^{(i)}|$. Thus, the factoring process must eventually stop; we denote the final factorization of w by $w^{(N)}$. We note that $w^{(N)}$ satisfies properties (5) – (7), and furthermore, $w^{(N)}$ satisfies property (5a) or (5b).

We now describe a set X of vertices in I_w . X is carefully defined so that it induces a set of components of I_w that is closed under transitivity of pairs and whose unique longest component is C . To define X , we will be using the final factorization $w^{(N)}$ of w , which has all the properties we need. The exact definition of X depends on whether $w^{(N)}$ satisfies property (5a) or (5b).

Let A denote the set of pairs of distinct subexpression-equivalent paths in $w^{(N)}$ of maximal length. Each pair of paths in A corresponds to a pair of distinct walks in Γ_w , say Y_0, Y_1, \dots, Y_r and Z_0, Z_1, \dots, Z_r , that trace out the same word; these two walks define a set of vertices in I_w , namely $\{Y_0, Z_0\}$,

$\{Y_1, Z_1\}, \dots, \{Y_r, Z_r\}\}$. If $w^{(N)}$ satisfies property (5a), we let X equal the union of all the sets of vertices obtained from pairs of paths in A that start with $u^{(N)}$. If $w^{(N)}$ satisfies property (5b), we let X equal the union of all the sets of vertices obtained from pairs of paths in A .

Claim 2. *The vertices in X induce a set of components of I_w that contains C and is closed under transitivity of pairs. Also, C is the (unique) longest component induced by X .*

Proof of Claim 2. We first suppose that $w^{(N)}$ satisfies property (5a). Consider any pair of paths in A that start with $u^{(N)}$. We claim that one of the paths ends right before a $b^{(N)}$ while the other path ends right before a $v^{(N)}$. This can be seen by noting that the pair of paths start with $u^{(N)}$, no occurrence of $u^{(N)}$ is cyclically followed by $u^{(N)}$, and any occurrence of $v^{(N)}$ or $b^{(N)}$ is always cyclically followed by $u^{(N)}$. Since $b_1^{(N)} \neq v_1^{(N)}$ (property (5a)), we see that the pair of walks in Γ_w corresponding to the pair of paths end at a pair of vertices that has degree 1 in I_w . Since $b_{|b^{(N)}|}^{(N)} \neq v_{|v^{(N)}|}^{(N)}$ and no occurrence of $u^{(N)}$ is cyclically preceded by $u^{(N)}$ (property (5a)), if the pair of walks in Γ_w start at a pair of vertices that has degree 2 in I_w , then the pair of initial $u^{(N)}$'s is cyclically preceded by a pair of $b^{(N)}$'s or a pair of $v^{(N)}$'s, both of which are cyclically preceded by a pair of $u^{(N)}$'s (property (6)). This shows that X induces a set of components of I_w ; furthermore, for every component induced by X , there exists a pair of paths in A starting with $u^{(N)}$ that produces exactly the vertices of the component. The case where $w^{(N)}$ satisfies property (5b) can be handled similarly.

We now show that X is closed under transitivity of pairs. We first suppose that $w^{(N)}$ satisfies property (5a). Let $\{U, V\}$ and $\{V, W\}$ be distinct vertices in X . Let D (respectively, D') be the component containing $\{U, V\}$ (respectively, $\{V, W\}$). Each component induced by X corresponds to some pair of paths in A that start with $u^{(N)}$. Thus, each component has a starting vertex and an ending vertex.

Suppose that $\{U, V\}$ is the starting vertex of D . Then, we note that $\{U, V\}$ corresponds to a pair of locations in $w^{(N)}$, each of which is cyclically in between two symbols (as opposed to being within a symbol). Since D corresponds to some pair of paths in A that start with $u^{(N)}$, the pair of symbols cyclically after (the pair of locations of) $\{U, V\}$ in $w^{(N)}$ are both $u^{(N)}$. Since D' contains $\{V, W\}$ and corresponds to some pair of paths in A , and since V corresponds to some location in $w^{(N)}$ that is cyclically in between two symbols, W also corresponds to some location in $w^{(N)}$ that is cyclically in between two symbols.

We claim that the symbol cyclically after (the location of) W in $w^{(N)}$ is

also a $u^{(N)}$. To see this, we note that the pair of paths corresponding to D' is in A and starts with $u^{(N)}$, so the pair of paths must end right before a $b^{(N)}$ and a $v^{(N)}$ (shown earlier in this proof of this claim). Thus, the pair of paths cannot end at the pair of symbols cyclically before $\{V, W\}$ since the symbol cyclically after V is a $u^{(N)}$, so the pair of paths must go past this pair of symbols. Thus, the symbol cyclically after W matches the symbol cyclically after V . Thus, the symbol cyclically after W in $w^{(N)}$ is also a $u^{(N)}$, so $\{U, W\}$ is in X . A similar argument shows that if $\{V, W\}$ is the starting vertex of D' , then $\{U, W\}$ is in X .

Now, suppose that both $\{U, V\}$ and $\{V, W\}$ are not starting vertices. Also, suppose that for every vertex $\{U', V'\}$ in D that is between the starting vertex and the vertex preceding $\{U, V\}$ (inclusive), if $\{U', V'\}$ (nontrivially) intersects any other vertex in X , then their symmetric difference is still in X . Then, the symmetric difference of the pair of vertices preceding $\{U, V\}$ and $\{V, W\}$ is in X , and we note that this vertex is a neighbor of $\{U, W\}$. Thus, $\{U, W\}$ is in X . The case where $w^{(N)}$ satisfies property (5b) can be handled similarly.

Since $w^{(N)}$ satisfies property (7), C is the longest component induced by X . □

This completes the proof of the theorem. □

5 Discussion of the $\phi(w) = \beta(w)$ Conjecture

The main conjecture we want to show is that $\phi(w) = \beta(w)$ for every word $w \in \Sigma^*$. We can rewrite this conjecture in the following form:

Conjecture 1. *For every word $w \in \Sigma^*$ and every non-negative integer i , we have $\phi(w) = i$ iff $\beta(w) = i$.*

We note that Conjecture 1 automatically implies that $\phi(w) = \infty$ iff $\beta(w) = \infty$. In [LP09], Linial and Puder implicitly make the following conjecture:

Conjecture 2. *For every word $w \in \Sigma^*$ and every non-negative integer i , if $\beta(w) \geq i$, then*

$$a_0(w) = \dots = a_{i-1}(w) = 0 \quad \text{and} \\ a_i(w) = |\{\Gamma \in \mathcal{Q}_w : \chi(\Gamma) = i \text{ and } \Gamma \text{ has type } B\}|.$$

It is easy to see that if Conjecture 2 is true, then for every word $w \in \Sigma^*$, we have $\phi(w) = \beta(w)$, and if $\beta(w) < \infty$, we have

$$\Phi_w(n) = \frac{|\{\Gamma \in \mathcal{Q}_w : \chi(\Gamma) = \beta(w) \text{ and } \Gamma \text{ has type } B\}|}{n^{\beta(w)}} + O\left(\frac{1}{n^{\beta(w)+1}}\right) \quad \text{and}$$

$$\sum_{\substack{\Gamma \in \mathcal{Q}_w \\ \Gamma \text{ has type A}}} P(\Gamma) = \Phi_w(n) + \frac{1}{n} - \sum_{\substack{\Gamma \in \mathcal{Q}_w \\ \Gamma \text{ has type B}}} P(\Gamma) = \frac{1}{n} + O\left(\frac{1}{n^{\beta(w)+1}}\right).$$

Recall that we have

$$\Phi_w(n) = -\frac{1}{n} + \sum_{\substack{\Gamma \in \mathcal{Q}_w \\ \Gamma \text{ has type A}}} P(\Gamma) + \sum_{\substack{\Gamma \in \mathcal{Q}_w \\ \Gamma \text{ has type B}}} P(\Gamma) = \sum_{i=0}^{\infty} a_i(w) \frac{1}{n^i}.$$

Conjecture 2 states that if \mathcal{Q}_w does not have any type-B quotients with characteristic less than i , then the contributions from the $-\frac{1}{n}$ term and the type-A quotient terms (in $\Phi_w(n)$ above) to the coefficients $a_0(w), \dots, a_i(w)$ cancel each other exactly so that $a_0(w) = \dots = a_{i-1}(w) = 0$ and $a_i(w) = |\{\Gamma \in \mathcal{Q}_w : \chi(\Gamma) = i \text{ and } \Gamma \text{ has type B}\}|$. In order to explain this cancellation, it may be tempting to conjecture that $\sum_{\substack{\Gamma \in \mathcal{Q}_w \\ \Gamma \text{ has type A}}} P(\Gamma) = \frac{1}{n}$ for every word $w \in \Sigma^*$ so that $\Phi_w(n) = \sum_{\substack{\Gamma \in \mathcal{Q}_w \\ \Gamma \text{ has type B}}} P(\Gamma)$ and Conjecture 2 holds. Or, it may be tempting to conjecture that $\sum_{\substack{\Gamma \in \mathcal{Q}_w \\ \Gamma \text{ has type A} \\ \chi(\Gamma) \leq i}} P(\Gamma) = \frac{1}{n}$ for every word $w \in \Sigma^*$ so that $\Phi_w(n) = \sum_{\substack{\Gamma \in \mathcal{Q}_w \\ \Gamma \text{ has type A} \\ \chi(\Gamma) > i}} P(\Gamma) + \sum_{\substack{\Gamma \in \mathcal{Q}_w \\ \Gamma \text{ has type B}}} P(\Gamma)$ and Conjecture 2 holds.

Neither of these conjectures are true; $w = g_1 g_1 g_2 g_2$ and $w = g_1 g_1 g_1 g_1$ are counterexamples showing that these two conjectures are false, respectively.

We know that $\beta(w)$, Φ_w , and the coefficients $a_0(w), a_1(w), \dots$ in the power series of Φ_w are all invariant under cyclic shift and reduction of w . Furthermore, it is implicitly shown in [LP09] that $|\{\Gamma \in \mathcal{Q}_w : \chi(\Gamma) = \beta(w) \text{ and } \Gamma \text{ has type B}\}|$ is also invariant under cyclic shift and reduction of w . Thus, when attempting to prove Conjecture 2, one can assume that the word $w \in \Sigma^*$ is cyclically reduced.

Recall that we have

$$\Phi_w(n) = -\frac{1}{n} + \sum_{\Gamma \in \mathcal{Q}_w} \left(\frac{1}{n}\right)^{\chi(\Gamma)} \cdot \prod_{l=1}^{v_{\Gamma}-1} \left(1 - \frac{l}{n}\right) \cdot \prod_{j=1}^k \prod_{l=1}^{e_{\Gamma}^j-1} \left(1 + \frac{l}{n} + \frac{l^2}{n^2} + \dots\right)$$

We can easily write explicit formulas for the first few coefficients of the power series $\sum_{i=0}^{\infty} a_i(w) \frac{1}{n^i}$ for Φ_w . One can easily verify the following:

$$a_0(w) = \begin{cases} 0 & \text{if } \Gamma_w \text{ has type A} \\ 1 & \text{if } \Gamma_w \text{ has type B} \end{cases}$$

$$a_1(w) = \begin{cases} |\{\Gamma \in \mathcal{Q}_w : \chi(\Gamma) = 1 \text{ and } \Gamma \text{ has type B}\}| & \text{if } \Gamma_w \text{ has type A} \\ -1 & \text{if } \Gamma_w \text{ has type B} \end{cases}$$

If Γ_w has type B (i.e., w reduces to the empty word), then $\Phi_w(n) = 1 - \frac{1}{n}$, so $a_2(w) = a_3(w) = \dots = 0$. Thus, assuming that Γ_w has type A, one can easily verify the following:

$$\begin{aligned} a_2(w) = & - \sum_{\substack{\Gamma \in \mathcal{Q}_w \\ \chi(\Gamma)=1}} \left(\binom{v_\Gamma}{2} - \sum_{j=1}^k \binom{e_\Gamma^j}{2} \right) + |\{\Gamma \in \mathcal{Q}_w : \chi(\Gamma) = 2\}| \\ a_3(w) = & \sum_{\substack{\Gamma \in \mathcal{Q}_w \\ \chi(\Gamma)=1}} \left(\sum_{i=1}^{v_\Gamma-1} \sum_{j=i+1}^{v_\Gamma-1} ij - \sum_{i=1}^{v_\Gamma-1} \sum_{j=1}^k \sum_{l=1}^{e_\Gamma^j-1} il + \sum_{j=1}^k \sum_{l=1}^{e_\Gamma^j-1} \sum_{l'=l+1}^{e_\Gamma^j-1} ll' + \sum_{j=1}^k \sum_{l=1}^{e_\Gamma^j-1} l^2 \right. \\ & \left. + \sum_{j=1}^k \sum_{j'=j+1}^k \sum_{l=1}^{e_\Gamma^j-1} \sum_{l'=1}^{e_{\Gamma'}^{j'}-1} ll' \right) - \sum_{\substack{\Gamma \in \mathcal{Q}_w \\ \chi(\Gamma)=2}} \left(\binom{v_\Gamma}{2} - \sum_{j=1}^k \binom{e_\Gamma^j}{2} \right) + |\{\Gamma \in \mathcal{Q}_w : \chi(\Gamma) = 3\}| \end{aligned}$$

There may be other ways of writing the contribution from characteristic 1 quotients to $a_3(w)$ that are more useful and meaningful. From the formulas for $a_0(w)$ and $a_1(w)$, we see that Conjecture 2 holds for $i = 0, 1$. In our proof of $\phi(w) = 2$ iff $\beta(w) = 2$, we have actually shown that Conjecture 2 holds for $i = 2$. It is easy to show that for every non-negative integer j , if Conjecture 1 holds for every $i < j$ and Conjecture 2 holds for $i = j$, then Conjecture 1 also holds for $i = j$. Using this fact, it is also easy to show (by induction on j) that for every non-negative integer j , if Conjecture 2 holds for every $i \leq j$, then Conjecture 1 also holds for every $i \leq j$.

We, as well as Linial and Puder, believe that Conjecture 2 is true, and this conjecture implies that $\phi(w) = \beta(w)$ for every word $w \in \Sigma^*$. We now know that Conjecture 2 is true for $i = 0, 1, 2$, and Linial and Puder have carried out extensive numerical simulations to test Conjecture 2 for $i = 3, 4$ without any failure (see [LP09]). We believe that a good strategy for proving the $\phi(w) = \beta(w)$ conjecture is to actually prove Conjecture 2.

Recall that in the proof of $\phi(w) = 2$ iff $\beta(w) = 2$, we showed there exists a bijection between the components of I_w and the quotients in \mathcal{Q}_w that have characteristic 2 and type A. Now, fix $w \in \Sigma^*$. From here on, we will assume that w is cyclically reduced, $\phi(w) \geq 3$, and $\beta(w) \geq 3$. Since $\beta(w) \geq 3$, we note that Γ_w is the only quotient (in \mathcal{Q}_w) of characteristic 1, and every

quotient of characteristic 2 has type A. We note that every type-A quotient of characteristic 3 can be obtained by joining two vertices of a quotient of characteristic 2 (and then performing a realization operation). I.e., for every type-A quotient $\Gamma \in \mathcal{Q}_w$ of characteristic 3, there exists a quotient $\Gamma' \in \mathcal{Q}_w$ of characteristic 2 and a component C of $I_{\Gamma'}$ such that $\text{rel}(\Gamma'/V(C)) = \Gamma$.

Looking at our formula for $a_3(w)$, we note that $\sum_{\substack{\Gamma \in \mathcal{Q}_w \\ \chi(\Gamma)=2}} \left(\binom{v_\Gamma}{2} - \sum_{j=1}^k \binom{e_\Gamma^j}{2} \right) = \sum_{\substack{\Gamma \in \mathcal{Q}_w \\ \chi(\Gamma)=2}} ((\# \text{ of components of } I_\Gamma) - (\# \text{ of independent cycles in } I_\Gamma))$. It is not hard to show that for every quotient $\Gamma \in \mathcal{Q}_w$ of characteristic 2, every component of I_Γ contains at most 1 cycle each. Thus, $\sum_{\substack{\Gamma \in \mathcal{Q}_w \\ \chi(\Gamma)=2}} \left(\binom{v_\Gamma}{2} - \sum_{j=1}^k \binom{e_\Gamma^j}{2} \right) = \sum_{\substack{\Gamma \in \mathcal{Q}_w \\ \chi(\Gamma)=2}} (\# \text{ of acyclic components of } I_\Gamma)$. Using the identification graph (of a quotient), it is not hard to show that for every quotient $\Gamma \in \mathcal{Q}_w$ of characteristic i , if C is a component of I_Γ that implies a component (possibly itself) containing a cycle, then $\text{rel}(\Gamma/V(C))$ has characteristic at most i . (This is actually true for any word $w \in \Sigma^*$.) Thus, for every type-A quotient Γ of characteristic 3, there exists a quotient Γ' of characteristic 2 and an *acyclic* component C of $I_{\Gamma'}$ such that $\text{rel}(\Gamma'/V(C)) = \Gamma$. Thus, $-\sum_{\substack{\Gamma \in \mathcal{Q}_w \\ \chi(\Gamma)=2}} \left(\binom{v_\Gamma}{2} - \sum_{j=1}^k \binom{e_\Gamma^j}{2} \right) + |\{\Gamma \in \mathcal{Q}_w : \chi(\Gamma) = 3 \text{ and } \Gamma \text{ has type A}\}| \leq 0$.

It may be true that for every quotient $\Gamma \in \mathcal{Q}_w$ of characteristic 2, if C is a component of I_Γ that does not imply a component containing a cycle, then $\text{rel}(\Gamma/V(C))$ has characteristic 3 (and thus type A as well). Assuming this is true, we have a surjective function from $\{C : C \text{ is a component of } I_\Gamma \text{ that does not imply a component containing a cycle, where } \Gamma \text{ is any quotient in } \mathcal{Q}_w \text{ of characteristic 2}\}$ to $\{\Gamma \in \mathcal{Q}_w : \chi(\Gamma) = 3 \text{ and } \Gamma \text{ has type A}\}$. This function is normally not injective, but it may be the case that this function is injective when restricted to $\{C : C \text{ is a component of } I_\Gamma \text{ that does not imply a component containing a cycle}\}$ for any single quotient Γ of characteristic 2. It may also be true that for every quotient $\Gamma \in \mathcal{Q}_w$ of characteristic 2, a component C of I_Γ does not imply a component containing a cycle iff C itself does not contain a cycle; this would simplify the above statements a bit.

In any case, we suspect that the contribution of Γ_w to $a_3(w)$ is always non-negative and exactly cancels the contribution of the quotients of characteristic 2 and the type-A quotients of characteristic 3. I.e, the contribution of Γ_w to $a_3(w)$ is $\sum_{\substack{\Gamma \in \mathcal{Q}_w \\ \chi(\Gamma)=2}} \left(\binom{v_\Gamma}{2} - \sum_{j=1}^k \binom{e_\Gamma^j}{2} \right) - |\{\Gamma \in \mathcal{Q}_w : \chi(\Gamma) = 3 \text{ and } \Gamma \text{ has type A}\}|$, so that $a_3(w) = |\{\Gamma \in \mathcal{Q}_w : \chi(\Gamma) = 3 \text{ and } \Gamma \text{ has type B}\}|$, as desired. The contribution of Γ_w to $a_3(w)$ seems to be compensating for the “lack of injectivity” of the above function.

It would be good to be able to express the contribution from quotients of characteristic 1 to $a_3(w)$ in a more useful and meaningful form. It would also be good to be able to formulate Conjecture 2 for $i = 3$ (and higher) as a simpler and more accessible combinatorial problem, like what was done for $i = 2$.

References

- [AL06] Alon Amit and Nathan Linial, *Random lifts of graphs II: Edge expansion*, Combinatorics Probability and Computing **15** (2006), no. 3, 317–332.
- [ALM02] Alon Amit, Nathan Linial, and Jirí Matousek, *Random lifts of graphs III: Independence and chromatic number*, Random Structures and Algorithms **20** (2002), no. 1, 1–22.
- [BS87] Andrei Broder and Eli Shamir, *On the second eigenvalue of random regular graphs*, Proceedings of the 28th Annual Symposium on Foundations of Computer Science, 1987, pp. 286–294.
- [Fri03] Joel Friedman, *Relative expanders or weakly relatively ramanujan graphs*, Duke Math. J. **118** (2003), 19–35.
- [Fri08] Joel Friedman, *A proof of Alon’s second eigenvalue conjecture and related problems*, Memoirs of the A.M.S **195** (2008).
- [Gre95] Y. Greenberg, *On the spectrum of graphs and their universal coverings*, Ph.D. thesis, Hebrew University of Jerusalem, 1995.
- [HLW06] Shlomo Hoory, Nathan Linial, and Avi Wigderson, *Expander graphs and their applications*, Bulletin of the American Mathematical Society **43** (2006), no. 4, 439–561.
- [LP09] Nati Linial and Doron Puder, *Words maps and spectra of random graph lifts*, accepted to Random Structures and Algorithms (2009).
- [LR05] Nathan Linial and Eyal Rozenman, *Random lifts of graphs: Perfect matchings*, Combinatorica **25** (2005), no. 4, 407–424.
- [Nil91] Alon Nilli, *On the second eigenvalue of a graph*, Discrete Mathematics **91** (1991), no. 2, 207–210.